



JA

RAW SEQUENCE LISTING **ERROR REPORT**

The Biotechnology Systems Branch of the Scientific and Technical Information Center (STIC) detected errors when processing the following computer readable form:

Application Serial Number: 10/069,353
Source: PCT10
Date Processed by STIC: 2/6/03

THE ATTACHED PRINTOUT EXPLAINS DETECTED ERRORS.

PLEASE FORWARD THIS INFORMATION TO THE APPLICANT BY EITHER:

- 1) INCLUDING A COPY OF THIS PRINTOUT IN YOUR NEXT COMMUNICATION TO THE APPLICANT, WITH A NOTICE TO COMPLY or,
- 2) TELEPHONING APPLICANT AND FAXING A COPY OF THIS PRINTOUT, WITH A NOTICE TO COMPLY

FOR CRF SUBMISSION QUESTIONS, PLEASE CONTACT MARK SPENCER, 703-308-4212.

FOR SEQUENCE RULES INTERPRETATION, PLEASE CONTACT ROBERT WAX, 703-308-4216.

PATENTIN 2.1 e-mail help: patin21help@uspto.gov or phone 703-306-4119 (R. Wax)

PATENTIN 3.0 e-mail help: patin3help@uspto.gov or phone 703-306-4119 (R. Wax)

TO REDUCE ERRORED SEQUENCE LISTINGS, PLEASE USE THE **CHECKER**
VERSION 3.1 PROGRAM, ACCESSIBLE THROUGH THE U.S. PATENT AND
TRADEMARK OFFICE WEBSITE. SEE BELOW FOR ADDRESS:

<http://www.uspto.gov/web/offices/pac/checker>

Applicants submitting genetic sequence information electronically on diskette or CD-Rom should be aware that there is a possibility that the disk/CD-Rom may have been affected by treatment given to all incoming mail.

Please consider using alternate methods of submission for the disk/CD-Rom or replacement disk/CD-Rom.

Any reply including a sequence listing in electronic form should NOT be sent to the 20231 zip code address for the United States Patent and Trademark Office, and instead should be sent via the following to the indicated addresses:

1. EFS-Bio (<<http://www.uspto.gov/ebc/efs/downloads/documents.htm>> , EFS Submission User Manual - ePAVE)
2. U.S. Postal Service: U.S. Patent and Trademark Office, Box Sequence, P.O. Box 2327, Arlington, VA 22202
3. Hand Carry directly to:
U.S. Patent and Trademark Office, Technology Center 1600, Reception Area, 7th Floor, Examiner Name,
Sequence Information, Crystal Mall One, 1911 South Clark Street, Arlington, VA 22202
Or
U.S. Patent and Trademark Office, Box Sequence, Customer Window, Lobby, Room 1B03, Crystal Plaza Two,
2011 South Clark Place, Arlington, VA 22202
4. Federal Express, United Parcel Service, or other delivery service to: U.S. Patent and Trademark Office,
Box Sequence, Room 1B03-Mailroom, Crystal Plaza Two, 2011 South Clark Place, Arlington, VA 22202

ERROR DETECTED**SUGGESTED CORRECTION**SERIAL NUMBER: 101069,353

ATTN: NEW RULES CASES: PLEASE DISREGARD ENGLISH "ALPHA" HEADERS, WHICH WERE INSERTED BY PTO SOFTWARE

- 1 Wrapped Nucleics
 Wrapped Aminos
The number/text at the end of each line "wrapped" down to the next line. This may occur if your file was retrieved in a word processor after creating it. Please adjust your right margin to .3; this will prevent "wrapping."
- 2 Invalid Line Length
The rules require that a line not exceed 72 characters in length. This includes white spaces.
- 3 Misaligned Amino
 Numbering
The numbering under each 5th amino acid is misaligned. Do not use tab codes between numbers; use space characters, instead.
- 4 Non-ASCII
The submitted file was not saved in ASCII(DOS) text, as required by the Sequence Rules. Please ensure your subsequent submission is saved in ASCII text.
- 5 Variable Length
Sequence(s) contain n's or Xaa's representing more than one residue. Per Sequence Rules, each n or Xaa can only represent a single residue. Please present the maximum number of each residue having variable length and indicate in the <220>-<223> section that some may be missing.
- 6 PatentIn 2.0
 "bug"
A "bug" in PatentIn version 2.0 has caused the <220>-<223> section to be missing from amino acid sequences(s) . Normally, PatentIn would automatically generate this section from the previously coded nucleic acid sequence. Please manually copy the relevant <220>-<223> section to the subsequent amino acid sequence. This applies to the mandatory <220>-<223> sections for Artificial or Unknown sequences.
- 7 Skipped Sequences
 (OLD RULES)
Sequence(s) missing. If intentional, please insert the following lines for each skipped sequence:
(2) INFORMATION FOR SEQ ID NO:X: (insert SEQ ID NO where "X" is shown)
(i) SEQUENCE CHARACTERISTICS: (Do not insert any subheadings under this heading)
(xi) SEQUENCE DESCRIPTION: SEQ ID NO:X: (insert SEQ ID NO where "X" is shown)
This sequence is intentionally skipped

Please also adjust the "(ii) NUMBER OF SEQUENCES:" response to include the skipped sequences.
- 8 Skipped Sequences
 (NEW RULES)
Sequence(s) missing. If intentional, please insert the following lines for each skipped sequence.
<210> sequence id number
<400> sequence id number
000
- 9 Use of n's or Xaa's
 (NEW RULES)
Use of n's and/or Xaa's have been detected in the Sequence Listing.
Per 1.823 of Sequence Rules, use of <220>-<223> is MANDATORY if n's or Xaa's are present.
In <220> to <223> section, please explain location of n or Xaa; and which residue n or Xaa represents.
- 10 Invalid <213>
 Response
Per 1.823 of Sequence Rules, the only valid <213> responses are: Unknown, Artificial Sequence, or scientific name (Genus/species). <220>-<223> section is required when <213> response is Unknown or is Artificial Sequence
- 11 Use of <220>
Sequence(s) missing the <220> "Feature" and associated numeric identifiers and responses.
Use of <220> to <223> is MANDATORY if <213> "Organism" response is "Artificial Sequence" or "Unknown." Please explain source of genetic material in <220> to <223> section.
(See "Federal Register," 06/01/1998, Vol. 63, No. 104, pp. 29631-32) (Sec. 1.823 of Sequence Rules)
- 12 PatentIn 2.0
 "bug"
Please do not use "Copy to Disk" function of PatentIn version 2.0. This causes a corrupted file, resulting in missing mandatory numeric identifiers and responses (as indicated on raw sequence listing). Instead, please use "File Manager" or any other manual means to copy file to floppy disk.
- 13 Misuse of n
n can only be used to represent a single nucleotide in a nucleic acid sequence. N is not used to represent any value not specifically a nucleotide.



PCT10

RAW SEQUENCE LISTING

PATENT APPLICATION: US/10/069,353

Does Not Comply

Corrector's Notice Needed

DATE: 02/06/2003
TIME: 14:27:25

Input Set : A:\LeA33955-US

Output Set: N:\CRF4\02052003\J069353.raw

3 <110> APPLICANT: Bayer Aktiengesellschaft
 5 <120> TITLE OF INVENTION: Nucleic acids coding for enzyme activities of
 6 spinosyn biosynthesis
 8 <130> FILE REFERENCE: Le A 33 955
 C--> 10 <140> CURRENT APPLICATION NUMBER: US/10/069,353
 C--> 11 <141> CURRENT FILING DATE: 2000-08-17
 13 <160> NUMBER OF SEQ ID NOS: 55
 15 <170> SOFTWARE: PatentIn Ver. 2.1

ERRORED SEQUENCES

7921 <210> SEQ ID NO: 42
 7922 <211> LENGTH: 2595
 7923 <212> TYPE: PRT
 7924 <213> ORGANISM: Saccharopolyspora spinosa
 7926 <400> SEQUENCE: 42
 7927 Met Ser Glu Ala Gly Asn Leu Ile Ala Val Ile Gly Leu Ser Cys Arg
 7928 1 5 10 15
 7930 Leu Pro Gln Ala Pro Asp Pro Ala Ser Phe Trp Arg Leu Leu Arg Thr
 7931 20 25 30
 7933 Gly Thr Asp Ala Ile Thr Thr Val Pro Glu Gly Arg Trp Gly Asp Pro
 7934 35 40 45
 7936 Leu Pro Gly Arg Asp Ala Pro Lys Gly Pro Glu Trp Gly Gly Phe Leu
 7937 50 55 60
 7939 Ala Asp Val Asp Cys Phe Asp Pro Glu Phe Phe Gly Ile Ser Pro Arg
 7940 65 70 75 80
 7942 Glu Ala Ala Ala Val Asp Pro Gln Gln Arg Leu Ala Leu Glu Leu Ala
 7943 85 90 95
 7945 Trp Glu Ala Leu Glu Asp Ala Gly Ile Pro Ala Gly Glu Leu Arg Gly
 7946 100 105 110
 7948 Thr Ala Ala Gly Val Phe Met Gly Ala Ile Ser Asp Asp Tyr Ala Ala
 7949 115 120 125
 7951 Leu Leu Arg Glu Ser Pro Pro Glu Val Ala Ala Gln Tyr Arg Leu Thr
 7952 130 135 140
 7954 Gly Thr His Arg Ser Leu Ile Ala Asn Arg Val Ser Tyr Val Leu Gly
 7955 145 150 155 160
 7957 Leu Arg Gly Pro Ser Leu Thr Val Asp Ser Gly Gln Ser Ser Ser Leu
 7958 165 170 175
 7960 Val Gly Val His Leu Ala Ser Glu Ser Leu Arg Arg Gly Glu Cys Thr
 7961 180 185 190
 7963 Ile Ala Leu Ala Gly Gly Val Asn Leu Asn Leu Ala Ala Glu Ser Asn
 7964 195 200 205

RAW SEQUENCE LISTING

DATE: 02/06/2003

PATENT APPLICATION: US/10/069,353

TIME: 14:27:25

Input Set : A:\LeA33955-US

Output Set: N:\CRF4\02052003\J069353.raw

*Insert first digit,
and move over 1 space,
ensuring that
the last digit
does not fall
underneath the next amino*

E--> 8113 Asp Ser Asp Ala Asp Glu Pro Ile Ala Val Ile Gly Met Gly Cys Arg
8114 995 1000 1005
8116 Phe Pro Gly Gly Val Ser Ser Pro Glu Glu Leu Trp Gln Leu Val Ala
8117 1010 1015 1020
8119 Ala Gly Arg Asp Val Val Ser Glu Phe Pro Ala Asp Arg Gly Trp Asp
E--> 8120 1025→ 1030 1035 1040
8122 Leu Glu Arg Ala Gly Thr Ser His Val Arg Ala Gly Gly Phe Leu His
8123 1045 1050 1055
8125 Gly Ala Pro Asp Phe Asp Pro Gly Phe Phe Arg Ile Ser Pro Arg Glu
8126 1060 1065 1070
8128 Ala Leu Ala Met Asp Pro Gln Gln Arg Leu Leu Leu Glu Ile Ala Trp
8129 1075 1080 1085
8131 Glu Ala Val Glu Arg Gly Gly Ile Asn Pro Gln His Leu His Gly Ser
8132 1090 1095 1100
8134 Gln Thr Gly Val Phe Val Gly Ala Thr Ser Leu Asp Tyr Gly Pro Arg
E--> 8135 1105→ 1110 1115 1120
8137 Leu His Glu Ala Ser Glu Glu Ala Ala Gly Tyr Val Leu Thr Gly Ser
8138 1125 1130 1135
8140 Thr Thr Ser Val Ala Ser Gly Arg Val Ala Tyr Ser Phe Gly Phe Glu
8141 1140 1145 1150
8143 Gly Pro Ala Val Thr Val Asp Thr Ala Cys Ser Ser Ser Leu Val Ala
8144 1155 1160 1165
8146 Leu His Leu Ala Cys Gln Ser Leu Arg Ser Gly Glu Cys Asp Leu Ala
8147 1170 1175 1180
8149 Leu Ala Gly Gly Val Thr Val Met Ala Thr Pro Gly Met Phe Val Glu
E--> 8150 1185→ 1190 1195 1200
8152 Phe Ser Arg Gln Arg Gly Leu Ala Pro Asp Gly Arg Cys Lys Ser Phe
8153 1205 1210 1215
8155 Ala Glu Ala Ala Asp Gly Thr Gly Trp Ser Glu Gly Ala Gly Leu Val
8156 1220 1225 1230
8158 Leu Leu Glu Arg Leu Ser Asp Ala Arg Arg Asn Gly His Glu Val Leu
8159 1235 1240 1245
8161 Ala Val Val Arg Gly Ser Ala Val Asn Gln Asp Gly Ala Ser Asn Gly
8162 1250 1255 1260
8164 Leu Thr Ala Pro Asn Gly Ser Ser Gln Gln Arg Val Ile Ala Gln Ala
E--> 8165 1265→ 1270 1275 1280
8167 Leu Ala Ser Ala Gly Leu Ser Val Ser Asp Val Asp Ala Val Glu Ala
8168 1285 1290 1295
8170 His Gly Thr Gly Thr Arg Leu Gly Asp Pro Ile Glu Ala Gln Ala Leu
8171 1300 1305 1310
8173 Ile Ala Thr Tyr Gly Gln Gly Arg Leu Pro Glu Arg Pro Leu Trp Leu
8174 1315 1320 1325
8176 Gly Ser Met Lys Ser Asn Ile Gly His Ala Gln Ala Ala Ala Gly Ile
8177 1330 1335 1340
8179 Ala Gly Val Met Lys Met Val Met Ala Met Arg His Gly Gln Leu Pro
E--> 8180 1345→ 1350 1355 1360
8182 Arg Thr Leu His Val Asp Glu Pro Thr Ser Gly Val Asp Trp Ser Ala
8183 1365 1370 1375
8185 Gly Thr Val Gln Leu Leu Thr Glu Asn Thr Pro Trp Pro Gly Ser Gly

of global
Sample errors
shown - exist

in Seq. 42, 44, 46,
48, 50

See also item 3

on error summary sheet.

VERIFICATION SUMMARY

PATENT APPLICATION: US/10/069,353

DATE: 02/06/2003

TIME: 14:27:27

Input Set : A:\LeA33955-US

Output Set: N:\CRF4\02052003\J069353.raw

L:10 M:270 C: Current Application Number differs, Replaced Application Number
L:11 M:271 C: Current Filing Date differs, Replaced Current Filing Date
L:8120 M:332 E: (32) Invalid/Missing Amino Acid Numbering, SEQ ID:42
M:332 Repeated in SeqNo=42
L:9169 M:332 E: (32) Invalid/Missing Amino Acid Numbering, SEQ ID:44
M:332 Repeated in SeqNo=44
L:10390 M:332 E: (32) Invalid/Missing Amino Acid Numbering, SEQ ID:46
M:332 Repeated in SeqNo=46
L:12239 M:332 E: (32) Invalid/Missing Amino Acid Numbering, SEQ ID:48
M:332 Repeated in SeqNo=48
L:14583 M:332 E: (32) Invalid/Missing Amino Acid Numbering, SEQ ID:50
M:332 Repeated in SeqNo=50